



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/514,413	11/15/2004	Nick Campbell	71109-014	6822

20277 7590 10/30/2007
MCDERMOTT WILL & EMERY LLP
600 13TH STREET, N.W.
WASHINGTON, DC 20005-3096

EXAMINER

SHAH, PARAS D

ART UNIT	PAPER NUMBER
----------	--------------

2626

MAIL DATE	DELIVERY MODE
-----------	---------------

10/30/2007

PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/514,413	Applicant(s) CAMPBELL ET AL.	
	Examiner Paras Shah	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 15 November 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-19 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-19 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 15 November 2004 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date <u>11/15/2004, 08/01/2005</u> . | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This communication is in response to the application filed on 11/15/2004. Claims 1-19 are pending and have been examined.

Priority

2. Receipt is acknowledged of papers submitted under 35 U.S.C. 119(a)-(d), which papers have been placed of record in the file.
3. It should be noted that since a priority benefit is claimed to a foreign application, reference in the specification identifying the foreign application must be made.

Information Disclosure Statement

4. The information disclosure statement (IDS) submitted on 11/14/2004 and 08/01/2005 is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.
5. The reference JP 2001-282277 was not considered by the examiner in the IDS filed on 11/15/2004 since no translation of the abstract was provided.

Specification

The following guidelines illustrate the preferred layout for the specification of a utility application. These guidelines are suggested for the applicant's use.

Arrangement of the Specification

As provided in 37 CFR 1.77(b), the specification of a utility application should include the following sections in order. Each of the lettered items should appear in

Art Unit: 2626

upper case, without underlining or bold type, as a section heading. If no text follows the section heading, the phrase "Not Applicable" should follow the section heading:

- (a) TITLE OF THE INVENTION.
- (b) CROSS-REFERENCE TO RELATED APPLICATIONS.
- (c) STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT.
- (d) THE NAMES OF THE PARTIES TO A JOINT RESEARCH AGREEMENT.
- (e) INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED ON A COMPACT DISC.
- (f) BACKGROUND OF THE INVENTION.
 - (1) Field of the Invention.
 - (2) Description of Related Art including information disclosed under 37 CFR 1.97 and 1.98.
- (g) BRIEF SUMMARY OF THE INVENTION.
- (h) BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING(S).
- (i) DETAILED DESCRIPTION OF THE INVENTION.
- (j) CLAIM OR CLAIMS (commencing on a separate sheet).
- (k) ABSTRACT OF THE DISCLOSURE (commencing on a separate sheet).
- (l) SEQUENCE LISTING (See MPEP § 2424 and 37 CFR 1.821-1.825. A "Sequence Listing" is required on paper if the application discloses a nucleotide or amino acid sequence as defined in 37 CFR 1.821(a) and if the required "Sequence Listing" is not submitted as an electronic document on compact disc).

6. The listing of the references on pages 23-24 should be removed from the Specification. An additional IDS can be submitted with the corresponding copies of the references for consideration by the Examiner.

7. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

Claim Rejections - 35 USC § 112

8. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Art Unit: 2626

9. Claims 1-19 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Please see below for the reasons of indefiniteness. Further, the limitations "a range that is generated stably" and "highly reliable portion" is not understandable as to what the applicant is seeking to claim. The mentioned limitations were interpreted to mean ranges where a syllabic nuclei s extracted and where a voiced region was determined.

10. The claims are generally narrative and indefinite, failing to conform with current U.S. practice. They appear to be a literal translation into English from a foreign document and are replete with grammatical and idiomatic errors. All of the above claims 1-19 should be corrected.

Claim Rejections - 35 USC § 103

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

12. Claims 1, 2, 4, 8, 9, 11, 14, 15, and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over in Lea *et al.* ("Algorithms for acoustic prosodic analysis") in view of Mermelstein ("Automatic segmentation of speech into syllabic units") in view of

Schmidbauer ("Syllable-based Segment-hypotheses Generation in Fluently spoken speech using Gross Articulatory features.).

As to claims 1, 8, and 14, Lea *et al.* teaches an apparatus for determining, based on speech waveform data, a portion reliably representing a feature of the speech waveform, comprising:

extracting means for calculating (see Figure 1, sonorant energy filter and energy calculation), from said data, distribution of an energy of a prescribed frequency range of said speech waveform on a time axis, and for extracting, among various syllables of said speech waveform, a range that is generated stably by a source of said speech waveform, based on the distribution and pitch of said speech waveform (see Figure 1) (e.g. From the figure, speech is input into the system. Then, energy calculation is done to determine the syllable units (voicing). Further, a stable range is determined from the boundary that is determined by pitch. (see page 42.7.1, right column, last paragraph-page 42.7.2, left column, lines 1-12));

estimating means for calculating (See Figure 1, energy calculation), from said data, distribution of spectrum of said speech waveform on the time axis, and for estimating, based on the spectral distribution on the time axis, a range of said speech waveform of which change is well controlled by said source (see Figure 2 and page 42.7.3, right column, 1st full paragraph) (e.g. In the cited section two types of methods are compared. A speech spectrum is obtained for both methods in order to determine the boundary for each syllable, which is well

controlled. The well-controlled portions is determined of the boundary extracted (e.g. reliable));

However, Lea does not specifically teach the minimum of a time distribution waveform.

Mermelstein does teach the use of a time distribution waveform for detecting local minimums (see Figure 1, and page 881, left column, sect. I, entire section) (e.g. The cited section uses a convex-hull to determine local minimum on a loudness versus time waveform.)

It would have been obvious to one of ordinary skilled in the art at the time the invention as made to have modified the separation of speech signal into quasi-syllables as taught by Lea with the use of a time-distribution waveform as taught by Mermelstein. The motivation to have combined the references involves the segmentation of speech into syllable units (see Abstract).

However, Lea in view of Mermelstein do not specifically teach the range being stably extracted by the source.

Schmidbauer does teach

means for determining that range which is extracted by said extracting means as the range generated stably by said source and of which speech waveform is estimated by said estimating means to be well controlled by said source, as a highly reliable portion of said speech waveform (page 10.9.3, left column, 3rd full paragraph-right column, line 18) (e.g. The cited portion discloses

the syllabic nuclei boundary estimate and then extraction of stable regions of the syllabic nuclei.)

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the determination of a reliable portion of a speech waveform as taught by Lea in view of Mermelstein with the inclusion of extracting stable regions as taught by Schmidbauer. The motivation to have combined the references involves the ability to do further processing including context specification and stress pattern of utterances (see page 10.9.1, left column, 3rd full paragraph).

As to claims 2, 9, and 15, Lea in view of Mermelstein in view of Schmidbauer teach all of the limitations as in claim 1 above.

Furthermore, Lea teaches wherein said extracting means includes voiced/unvoiced determining means for determining, based on said data, whether each segment of said speech waveform is a voiced segment or not (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. Voiced and unvoiced determination is made.) of said waveform of energy distribution of the prescribed frequency range of said speech waveform on the time axis (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. In the cited section a prescribed frequency range is used and dips of energy define minimums.),

Furthermore, Mermelstein teaches the means for separating said speech waveform into syllables at a local minimum (see page 881, right column, 1st full paragraph, and Figure 1) (e.g. The minimum of Figure 1 is used to determine and segment syllable.); and

Furthermore, Lea teaches the means for extracting that range of said speech waveform which includes, in each syllable, an energy peak in that syllable within the segment determined to be a voiced segment by said voiced/unvoiced determining means and in which the energy of the prescribed frequency range is not lower than a prescribed threshold value (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. A threshold is used to determine voiced and unvoiced segments. A frequency range for sonorant energy is defined and since dips are located it is seen intuitively that maximums will occur.)

As to claims 4,11, and 17 Lea in view of Mermelstein in view of Schmidbauer teach all of the limitations as in claim 1 above.

Furthermore, Lea wherein said determining means includes means for determining, as a highly reliable portion of said speech waveform, a range included in the range extracted by said extracting means, within the range of which change in speech waveform is estimated by said estimating means to be well controlled by said source (see Figure 2 and page 42.7.3, right column, 1st full paragraph) (e.g. Form the figure, the syllables are detected and a range in time is

specified as seen on the frames on the x-axis, "island of reliability") (e.g. It would have been obvious to extract the frames corresponding to the extracted syllable as defined by the timing in the Figure (e.g. Frames). Further, the use of a voice detector as denoted in Lea will provide a range for voicing compared to unvoiced segments.)

13. Claims 5, 6, 12, 18, and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lea in view of Mermelstein.

As to claims 5, 12 and 18, Lea teaches a quasi-syllabic nuclei extracting apparatus for separating a speech signal into quasi-syllables and extracting a nuclear portion of each quasi-syllable, comprising:

voiced/unvoiced determining means (see Figure 1, voicing decision) for determining whether each segment of the speech signal is voiced or not (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1, voicing decision) (e.g. Voiced and unvoiced determination is made.);

means for separating said speech signal into quasi-syllables (see Figure 1 syllabic nucleus detection) at a local minimum of time-distribution waveform of an energy of a prescribed frequency range of said speech signal (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. In the cited section a prescribed frequency range is used and dips of energy define minimums (The term quasi-syllable was interpreted to mean relating to a syllable.); and

means for extracting that range of said speech signal which includes energy peak in each quasi-syllable (see Figure 1, energy calculation and syllabic energy detector), determined by said voiced/unvoiced determining means to be a voiced segment and of which energy of the prescribed frequency range is not lower than a prescribed threshold value, as the nuclei of quasi-syllable (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. A threshold is used to determine voiced and unvoiced segments. A frequency range for sonorant energy is defined and since dips are located it is seen intuitively that maximums will occur. Both the syllabic nucleus detection and voicing decision are interconnected.).

However, Lea does not specifically teach the minimum of a time distribution waveform.

Mermelstein does teach the use of a time distribution waveform for detecting local minimums (see Figure 1, and page 881, left column, sect. I, entire section) (e.g. The cited section uses a convex-hull to determine local minimum on a loudness versus time waveform.).

It would have been obvious to one of ordinary skilled in the art at the time the invention as made to have modified the separation of speech signal into quasi-syllables as taught by Lea with the use of a time-distribution waveform as taught by Mermelstein. The motivation to have combined the references involves the ability segment of speech into syllable units (see Abstract) more effectively.

As to claims 6 and 19, Lea in view of Mermelstein teach all of the limitations as in claims 5 and 18, above.

Furthermore, Lea teaches wherein said extracting means includes means for extracting that range of said speech signal which includes an energy peak in each pseudo-syllable within the segment determined to be a voiced segment by said voiced/unvoiced determining means and in which the energy of said prescribed frequency range is not lower than a prescribed threshold value as the nuclei of quasi-syllable (see page 42.7.1, right column, sect. 2, 1st full paragraph, and Figure 1) (e.g. A threshold is used to determine voiced and unvoiced segments. A frequency range for sonorant energy is defined and since dips are located it is seen intuitively that maximums will occur.). Furthermore, Mermelstein teaches the use of determining the peak of the loudness function in order to determine the syllable boundary (see page 881, right column, 1st full paragraph and Figure 1).

Allowable Subject Matter

14. Claims 3 and 16 would be allowable if rewritten to overcome the rejection(s) under 35 U.S.C. 112, 2nd paragraph, set forth in this Office action and to include all of the limitations of the base claim and any intervening claims.
15. Claim 7 would be allowable if rewritten or amended to overcome the rejection(s) under 35 U.S.C. 112, 2nd paragraph, set forth in this Office action.

16. The following is a statement of reasons for the indication of allowable subject matter: None of the prior arts or combination thereof teach the limitations as recited in claims 3, 7, and 16 as that of "based on an output from said linear predicting means, distribution on the time axis of local variance of spectral change" and "... means for estimating, based on both ... first calculating means and ... second calculating means". Most of the prior arts disclose the inclusion of the first calculating means.

Conclusion

17. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Lin *et al.* (US 4,802,223) is cited to disclose analyzing spoken input for syllables. Mekata (US 5,479,560) is cited to disclose a formant detection device. Hosom *et al.* (US 5,577,160) is cited to disclose an LPC analysis on speech waveform for extracting glottal parameters and formant parameters. Singhal (US 5,675,705) is cited to disclose syllable recognition. Kobayashi (US 7,035,798) is cited to disclose speech section detection using LPC and spectrum analysis. Yamoto *et al.* (US 7,231,346) is cited to disclose speech section detection for detecting speech sounds. Brandman (US 2003/0014245) is cited to disclose speech feature extraction system for speech recognition. Ealey *et al.* (US 2004/0133424) is cited to disclose the processing of speech signals for determining pitch and frequency.

The NPL document by Mercier *et al.* ("Automatic segmentation, Recognition of phonetic units and training in the KEAL speech recognition system") is cited to disclose

segmenting speech into linguistic units. *Lea et al.* ("A prosodically guided speech understanding strategy") is cited to disclose detecting boundaries from fall rise patterns of frequency contours.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Paras Shah whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:30a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571)272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

P.S.

10/23/2007


PATRICK N. EDOUARD
SUPERVISORY PATENT EXAMINER